

SURVEYING LOWER SECONDARY STUDENTS’ UNDERSTANDINGS OF ALGEBRA AND MULTIPLICATIVE REASONING: TO WHAT EXTENT DO PARTICULAR ERRORS AND INCORRECT STRATEGIES INDICATE MORE SOPHISTICATED UNDERSTANDINGS?

Jeremy Hodgen^{*}, Margaret Brown^{*}, Robert Coe^{**} and Dietmar Küchemann^{*}

^{*}King’s College London/^{**}University of Durham

jeremy.hodgen@kcl.ac.uk

In this paper, we will report on an analysis of student’s incorrect responses based on a large-scale assessment of algebra, ratio and decimals in lower secondary school in England. Our aim is to investigate the extent to which some errors or misconceptions may develop as students’ understanding increases and, hence, may be regarded as “better” errors that, although incorrect, indicate more sophisticated understanding. Using item response graphs for correct and incorrect responses, we indicate the potential for such analysis.

Algebra, Multiplicative reasoning, Errors, Misconceptions, Assessment.

INTRODUCTION

In this paper, we will report on an analysis of student’s incorrect responses based on a large-scale assessment of algebra, ratio and decimals in lower secondary school in England. Our aim is to investigate the extent to which some errors or misconceptions may develop as students’ understanding increases and, hence, may be regarded as “better” errors that, although incorrect, indicate more sophisticated understanding. This is important because, if some errors are indeed “better” than others, then it is likely that these may need different pedagogical approaches.

We note that in mathematics education there is a developing interest in the use of psychometric approaches which is partly driven by a dissatisfaction with existing approaches to the modeling and analysis of such data, particularly related to surveys of teachers’ mathematical knowledge (e.g., Izsák, Orrill, Cohen, & Brown, 2010) and partly by an interest in providing better advice to teachers on the ways in which they can help students (e.g., Nguyen, Rupp, Confrey, & Maloney). This work is largely focused on correct / incorrect responses, or involve multiple choice items (e.g., Andrich & Styles). Extending psychometric techniques to the analysis of incorrect responses has some difficulties, in particular the extent to which psychometric models can accommodate the idiosyncratic aspects of students’ learning (Denvir & Brown, 1986). We are currently using a range of psychometric techniques, although this analysis is not yet complete. In this paper,

we focus on a range of key items. Our aim is to indicate the potential for this kind of analysis.

Background and Methodology

Increasing Competence and Confidence in Algebra and Multiplicative Structures (ICCAMS) is a 4-year research project funded by the Economic and Social Research Council in the UK. Phase 1 of the project consisted of a cross-sectional survey of 11-14 years olds' understandings of algebra and multiplicative reasoning, and their attitudes to mathematics. The survey used tests of Algebra, Ratio and Decimals developed in the Concepts in Secondary Mathematics and Science (CSMS) study (Hart et al., 1981). These tests were first administered in 1976 and 1977 to a representative sample of lower secondary students in England. The aim of Phase 1 of ICCAMS was to examine how students' understandings have changed since the 1970s (Hodgen, Brown, Küchemann, & Coe, 2011) and to conduct a detailed analysis of current students' understandings.

Phase 2 of the study consisted of a collaborative research study with a group of teachers examining how formative assessment, and other forms of mathematical pedagogy supported by the results of educational research, could be used to improve attainment and attitudes in algebra and multiplicative reasoning (Brown, Hodgen & Küchemann, 2012). In Phase 3, recently completed, we examined how the work could be implemented on a larger scale with a larger group of schools and teachers. In this paper, we only report on Phase 1 of the research.

Over two summers in 2008 and 2009, the three tests (Algebra, Ratio, Decimals) were administered to a sample of approximately 8000 students across lower secondary (aged 11-14) from 19 schools randomly selected within strata of performance on the Middle Years Information System (MidYIS) database. MidYIS is a value added reporting system provided by Durham University, which is widely used across England. We have shown that this sample is approximately representative of the English population and how item facilities and other statistics can be estimated within acceptable levels (Coe, 2011). Each student took two of the three tests so as to provide comparative information between tests but not to overload students. The numbers of students in each year-group taking each test is therefore roughly two-thirds of the total number of students involved in that age group. (See Table 1.)

Table 1: Sample sizes for ICCAMS Tests 2008/9

Algebra	Ratio	Decimals
5470	5345	5485

The test items were developed on the basis of extensive diagnostic interviews with students with the aim of using “problems which were recognisably connected to the mathematics curriculum but which would require the child to use methods which were not obviously ‘rules’ ” (Hart & Johnson, 1983). For each test, the test items range from the basic to the

sophisticated allowing broad stages of attainment in each topic to be reported. The algebra test focuses on generalised arithmetic and variable. The ratio test focuses on the extent to which situations involving ratio are understood as multiplicative. The decimals test focuses on measurement and multiplicative aspects of decimal number.

Although the tests were designed in the mid-1970s the content has been shown by analysis of curriculum documents to be central to the current national curriculum and national assessment in England. Indeed, the CSMS results were influential in the construction of the English National Curriculum in 1989 (Brown, 1996). Piloting (conducted in early 2008) indicated that only minor updating of language or context for a very small number of items was required which would be unlikely to significantly affect their difficulty. The great majority of the items in the Decimals and Algebra tests were in any case mainly expressed in mathematical symbols. We have conducted a Rasch analysis using the Winsteps package. This indicates that each test can be considered as unidimensional and that all the items have an acceptable level of fit.

Items are largely in open-response format. Students' responses were captured on a database then coded. The coding aimed to capture responses that we expected from research and theory in addition to responses with a frequency of greater than 1%. This process has enabled us to capture the range of correct and incorrect responses and to identify similar response types across different items. In this paper, we focus on item response graphs that indicate the proportion of students attaining a particular total score who give the various coded responses. (See, for example, Figure 1.) This allows us to examine how the proportion of incorrect responses varies across the attainment range.

The comparison with the 1970s has been reported elsewhere (Brown, Küchemann, & Hodgen, 2010; Hodgen, Brown, Küchemann, & Coe, 2011; Hodgen, Küchemann, Brown, & Coe, 2009, 2010). Briefly, in all the tests, results were the same or slightly worse than in the 1970s, with more children now scoring very few marks, and fewer scoring very high marks. A comparison of student errors indicates broadly similar patterns of errors, although it should be noted that the coding of errors in the recent survey is more detailed than was practicable in the 1970s.

UNDERSTANDING STUDENT ERRORS

There has been a great deal of research indicating that many errors result from students trying to make sense of school mathematics (e.g., Brown & Van Lehan, 1982) and there is a vast literature on students' misconceptions (e.g., Booth, 1984; Ryan & Williams, 2007).

Multiplication Makes Bigger, Division Makes Smaller

Greer (1994) identified Multiplication Makes Bigger, Division Makes Smaller (MMBDMS) as a common misconception, where students generalise results involving whole number multiplication to conclude that multiplication always results in a larger number. A related error occurs very frequently on one of the items on the Algebra Test. Students are asked which is bigger, $2n$ or $n+2$, and to explain their response. Across the ICCAMS sample of lower secondary students, by far the most popular response, given by 48% of students, was of the sort " $2n$, because multiplication makes things bigger" (MMB). Only 1% give a

correct response of “It depends”, together with an explanation showing some awareness that $2n$ can be smaller, equal to, or larger than $n+2$, depending on the value of n . However, interestingly, the MMB response does appear to be associated with quite high levels of understanding. Figure 1 plots the proportion of students giving a particular response for each particular overall score on the test. This indicates that the MMB response occurs most frequently amongst students with total scores in the 20-45 range and occurs much less frequently amongst students with lower scores. Further, Figure 1 shows that demonstrating an awareness of the contingent nature of MMB is associated with very high test scores.

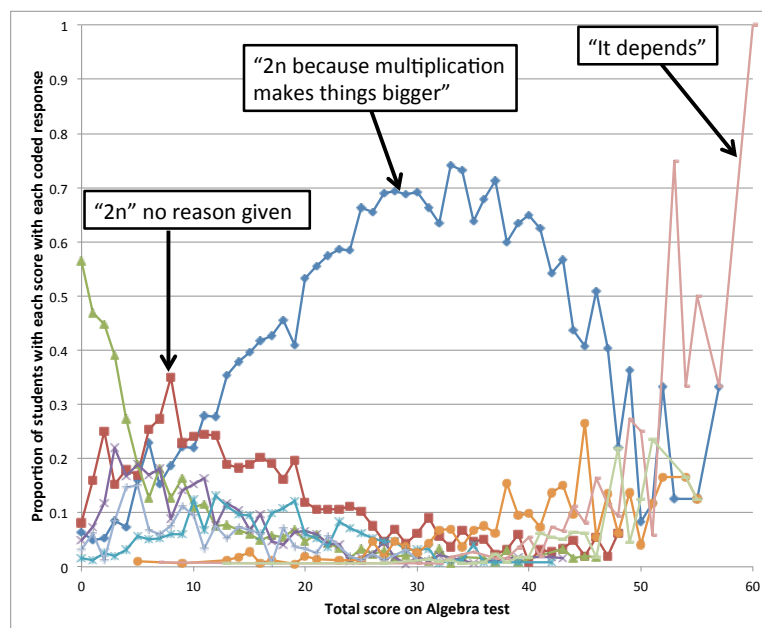


Figure 1. “Which is larger, $2n$ or $n + 2$? Explain why”: The proportion of students giving most frequent coded responses for each overall score on the Algebra test

Using an addition strategy to solve ratio problems

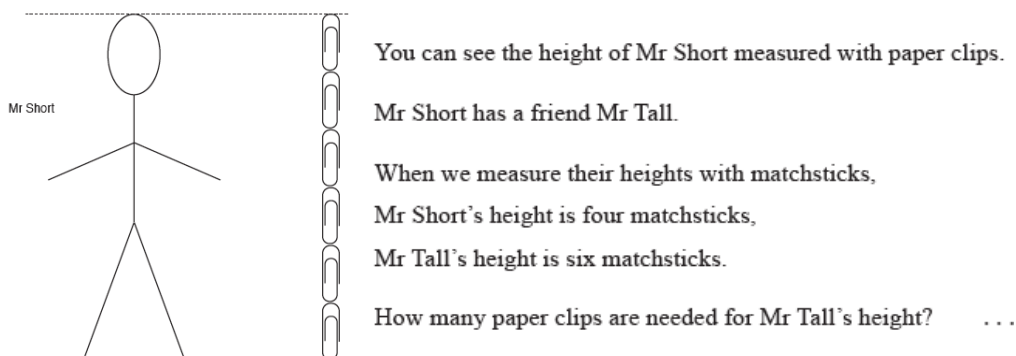


Figure 2: The Mr Short and Mr Tall problem

The relationship between student understandings and their errors and misconceptions have been analysed quantitatively in the past. Hart (1984; 1980), for example, investigated the understandings of “adders”, students who frequently use an inappropriate additive strategy to solve ratio problems. This ‘addition strategy’ was observed in the current study. The ratio test includes a version of the Mr Short & Mr Tall problem developed by Karplus and colleagues (Karplus & Petersen, 1970). (See Figure 2.)

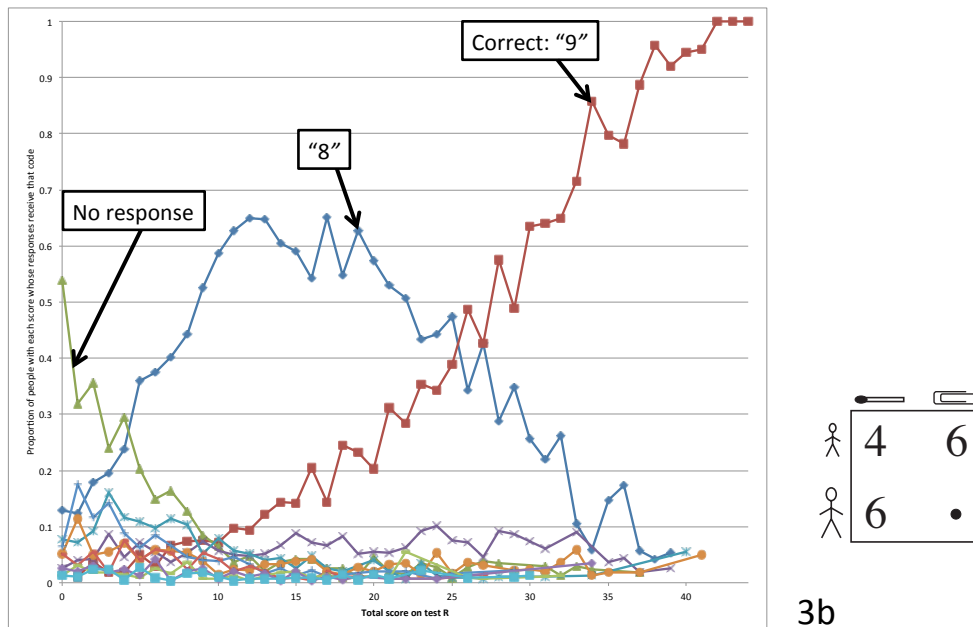


Figure 3a and b: The Mr Short and Mr Tall problem. (3a) The proportion of students giving most frequent coded responses for each overall score on the Ratio test. (3b) A ratio table indicating the structure of the problem.

We found, as Hart had done previously, that the most common response given by lower secondary school students is to state that Mr Tall needs 8 paperclips (rather than 9). This can be derived additively by arguing either that “Mr Short needs 2 more paper clips than matchsticks, so Mr Tall will need 2 more paper clips than matchsticks”, or “Mr Tall needs two more matchsticks than Mr Short, so he will need 2 more paper clips than Mr Short”. As can be seen from Fig 3a, such additive responses are common amongst students who do quite well on the ratio test as a whole, but that the frequency falls off rapidly for scores below 9 and above 22. We would argue that an addition strategy response requires students to *coordinate* the information that they are given (for example by putting it in a ratio table as in Fig 3b), and that this is not trivial.

Procedural and syntactic responses

Several items appear to promote incorrect procedural responses where students draw on inappropriate (but partially correct) syntactic understandings without checking for semantic meanings (Oldenburg, 2011). Figure 4 below is for an item from the Decimals test, again for our combined sample of 11-14 year olds. Students are asked to “write eleven tenths as a decimal” and it can be seen from the graph that the item discriminates extremely well (since

the correct response 0.11 is given by most students who score very highly on the test and is rarely given by student who do not score very highly on the test). The graph also shows that 0.11 is a common response to this item. It is tempting to class this as a gross error, but the graph suggests otherwise, since this is a common response amongst high scoring students. Perhaps students who give this response demonstrate some understanding that tenths are represented by the first column to the right of the decimal point, but then are unaware of, or unable to resolve, the contradictions that flow from placing more than one numeral into this column.

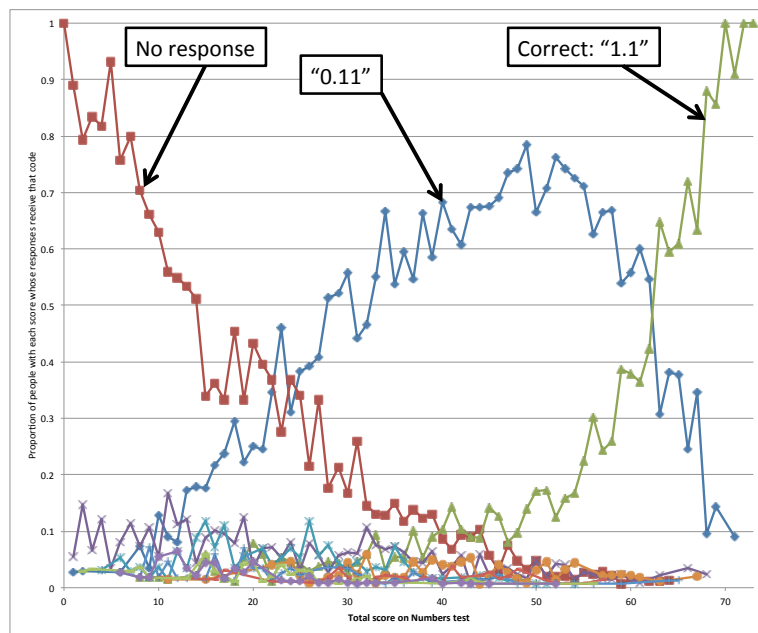


Figure 4. “Write eleven tenths as a decimal”: The proportion of students giving most frequent coded responses for each overall score on the Decimals test

Fig 5 shows three items 4d, 4e and 4f from the CSMS Algebra test. Though it might be argued that 4e and 4f are purely syntactical, they have very different facilities, suggesting that this is not the case.

n multiplied by 4 can be written as **$4n$** .
 Multiply each of these by 4:

8	$n + 5$	$3n$
.....

Figure 5: Items 4d, 4e and 4f from the CSMS Algebra test

In Fig 6 we have highlighted two partially correct responses, “ $n+20$ ” and “ $4n+5$ ”, and two correct responses, “ $4n+20$ ” and “ $4(n+5)$ ”. It can be seen from the graph that students giving the partially correct responses tend to score quite highly on the test as a whole, which can perhaps in part be explained by the fact that both responses require an “acceptance of lack

of closure” (Collis 1978). Further, the finding that “ $4n+5$ ” tends to be given by students who score markedly more highly than those giving “ $n+20$ ” suggests that some students might intend “ $4n+5$ ” to mean “4 times ($n+5$)” but lack familiarity with bracket notation or an awareness of the ambiguity of their expression. It also suggest that some students who give “ $n+20$ ” are simply combining the elements that are familiar to them ($4 \times 5 = 20$), and in effect leaving the algebraic symbol safely to one side. Regarding the two correct responses, the fact that fewer (high scoring) students give the correct response $4(n+5)$ than the correct response $4n+20$ might again in part be due to a lack of familiarity with bracket notation for some students in our sample, and perhaps also through a belief by some that an expression involving brackets has not been fully worked out.

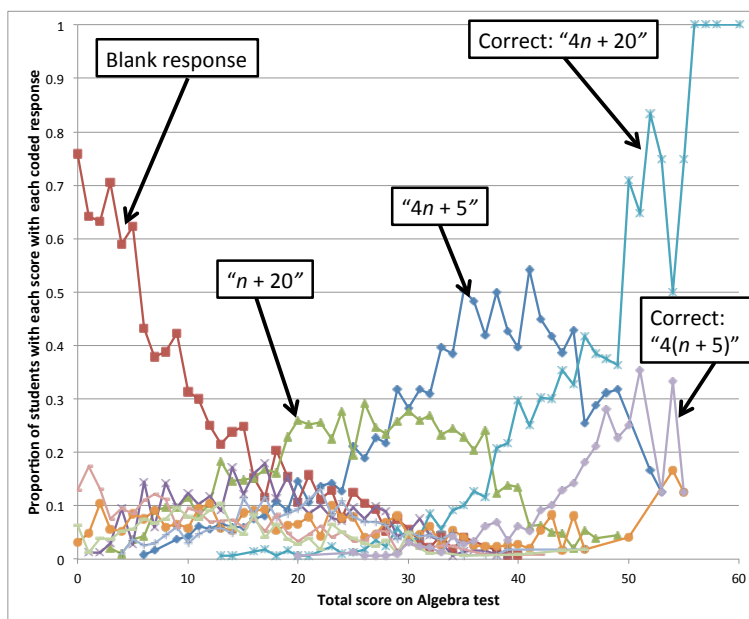


Figure 6: Multiply $n+5$ by 4. The proportion of students giving most frequent coded responses for each overall score on the Algebra test

Fig 7 is for item 4f. The most interesting feature here is the fact that the set of points for the correct response “ $12n$ ” show only a gradual rise in relative frequency for an increase in total score, in marked contrast to all the other graphs considered in this paper, where the set of points for the correct response shows a sudden and steep increase in relative frequency. In other words, this item discriminates less well. This is perhaps in part because the correct response can be achieved using a similar, partially correct strategy as for the response “ $n+20$ ” for item 4e, namely combining the familiar elements ($4 \times 3 = 12$) and leaving the symbol n as it was, adjacent to the numeral. Further evidence for this comes from Table 2, where it can be seen that students who give the correct response “ $12n$ ” on item 4f, tend to give the relatively low-level response “ $n+20$ ” to 4e.

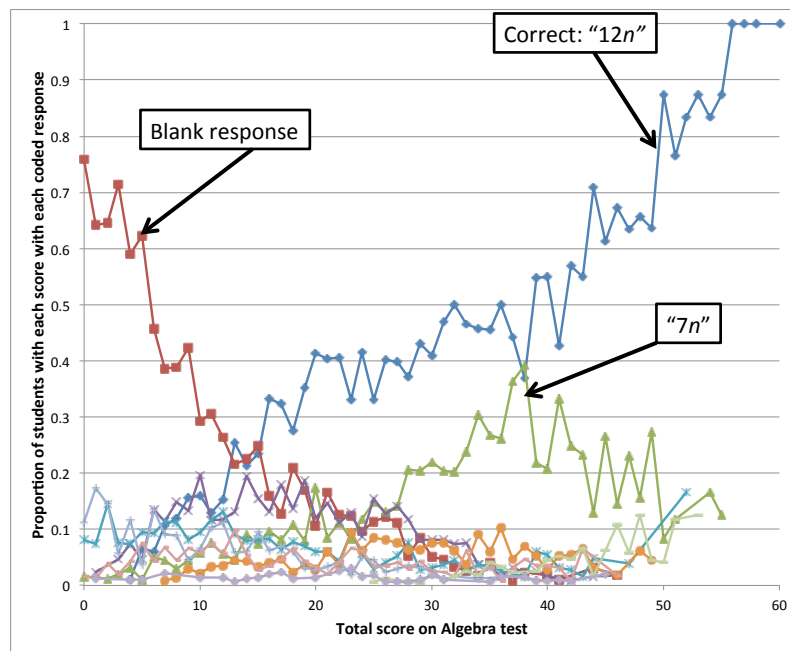


Figure 7: Multiply $3n$ by 4. The proportion of students giving most frequent coded responses for each overall score on the algebra test

Table 2: The inter-relationship between coded responses for two items from the stem “Multiply each of these by 4”: $3n$ and $n+5$. Note: Not all coded responses are shown, so the columns and rows do not sum to the totals shown.

		Multiply $3n$ by 4			Totals
		$12n$	$7n$	12	
Multiply $n+5$ by 4	$4n+5$	7%	12%	0%	25%
	$n+20$	15%	1%	1%	20%
	20	1%	0%	6%	8%
	$4n+20$	7%	1%	0%	9%
	Totals	41%	18%	11%	100%

Conclusion

We have outlined an approach to analysing performance data where a range of responses to a given test item are related to students’ scores on the test as a whole. Regarding students’ errors, this can, first of all, alert us to the fact that these errors, rather than arising merely from carelessness or ignorance, can be the result of serious mathematical thinking, albeit based on conceptualisations that may be limited in some way. In turn this can prompt us to

consider more deeply the nature of these conceptualisations and how we can help students continually to revise them (Nunes & Bryant 1996).

Currently our analysis using psychometric models is ongoing. We are investigating the use of a variety of models, including the partial credit Rasch model (e.g., Andrich & Styles, 2008), although the modelling has proved technically more difficult than we anticipated. A significant technical difficulty is associated with the fact that relatively few items indicate such clear patterns (less than 10%) and collapsing the responses into an ordered structure has not always been possible. We expect to resolve this difficulty shortly and to be able to report this analysis at ICME.

Acknowledgement

We are grateful to the Economic and Social Research Council in the UK for funding this study (RES-179-25-0009).

References

- Andrich, D., & Styles, I. M. (2008). *Identifying distractors which justify partial credit in multiple choice items: a routine application of a polytomous Rasch model hypothesis*. . Paper presented at the Third International Conference on Measurement in Health, Education, Psychology and Marketing: Developments with Rasch Models, Perth, Western Australia.
- Booth, L. (1984). *Algebra: Children's strategies and errors*. Windsor: NFER-NELSON.
- Brown, J. S., & Van Lehan, K. (1982). Towards a generative theory of 'bugs'. In T. P. Carpenter, J. M. Moser & T. A. Romberg (Eds.), *Addition and Subtraction: A cognitive perspective*. Hillsdale, NJ: Lawrence Erlbaum.
- Brown, M., Hodgen, J., & Küchemann, D. (2012). *Changing the Grade 7 curriculum in algebra and multiplicative thinking at classroom level in response to assessment data*. Paper to be presented at the 12th International Congress on Mathematical Education (ICME-12), Topic Study Group 32, Seoul, Korea.
- Brown, M., Küchemann, D. E., & Hodgen, J. (2010). The struggle to achieve multiplicative reasoning 11-14. In M. Joubert & P. Andrews (Eds.), *Proceedings of the Seventh British Congress of Mathematics Education (BCME7)* (Vol. 30, pp. 49-56). University of Manchester: BSRLM.
- Coe, R. (2011). *Representativeness of the ICCAMS sample: Technical report for ICCAMS (Increasing Competence and Confidence in Algebra and Multiplicative Structures)*. CEM Centre. Durham.
- Collis, K. F. (1978). Operational thinking in elementary mathematics. In J. A. Keats, K. F. Collis & G. S. Halford (Eds.), *Cognitive development: Research based on a Neo-Piagetian approach* (pp. 221-248). Chichester: John Wiley & Sons.
- Denvir, B., & Brown, M. (1986). Understanding number concepts in low attaining 7-9 year olds. Part I: Development of descriptive framework and diagnostic instruments. *Educational Studies in Mathematics*, 17, 15-36.

- Greer, B. (1994). Extending the Meaning of Multiplication and Division. In G. Harel & J. Confrey (Eds.), *The Development of Multiplicative Reasoning in the Learning of Mathematics* (pp. 61-85): SUNY Press.
- Hart, K. (1980). *Secondary school-children's understanding of ratio and proportion*. PhD thesis, Chelsea College, University of London. Hart, K. (1984). *Ratio: Children's strategies and errors*. Windsor: NFER-Nelson.
- Hart, K., Brown, M. L., Küchemann, D. E., Kerslake, D., Ruddock, G., & McCartney, M. (Eds.). (1981). *Children's understanding of mathematics: 11-16*. London: John Murray.
- Hart, K., & Johnson, D. C. (Eds.). (1983). *Secondary school children's understanding of mathematics: A report of the mathematics component of the concepts in secondary mathematics and science programme*. London: Centre for Science Education, Chelsea College.
- Hodgen, J., Brown, M., Küchemann, D., & Coe, R. (2011). *Why have educational standards changed so little over time: The case of school mathematics in England*. Paper presented at the British Educational Research Association (BERA) Annual Conference, Institute of Education, University of London.
- Hodgen, J., Küchemann, D., Brown, M., & Coe, R. (2009). Children's understandings of algebra 30 years on. *Research in Mathematics Education*, 11(2), 193-194.
- Hodgen, J., Küchemann, D., Brown, M., & Coe, R. (2010). Multiplicative reasoning, ratio and decimals: A 30 year comparison of lower secondary students' understandings. In M. F. Pinto & T. F. Kawaski (Eds.), *Proceedings of the 34th Conference of the International Group of the Psychology of Mathematics Education* (Vol. 3, pp. 89-96). Belo Horizonte, Brazil.
- Izsák, A., Orrill, C. H., Cohen, A. S., & Brown, R. E. (2010). Using the mixture Rasch model to assess middle grades teachers' reasoning about rational numbers. *Elementary School Journal*, 110(3), 279-300.
- Karplus, R., & Petersen, R. (1970). *Intellectual Development Beyond Elementary School 11. Ratio, a Survey: Science Curriculum Improvement Study*, Lawrence Hall of Science, University of California, Berkeley.
- Nguyen, K. H., Rupp, A. A., Confrey, J., & Maloney, A. P. (2012). *Testing the reorganization of the equipartitioning learning trajectory using Rasch Item Response theory modeling*. Paper presented at the 2012 Annual Conference of the American Educational Research Association (AERA), Vancouver.
- Nunes, T., & Bryant, P. (1996). *Children doing mathematics*. Oxford: Blackwell.
- Oldenburg, R. (2009). Structure of algebraic competencies. In V. Durand-Guerrier, S. Soury-Lavergne & F. Arzarello (Eds.), *Proceedings of the Sixth Congress of the European Society for Research in Mathematics Education (CERME 6)* (pp. 579-588). Lyon, France: Institut National De Recherche Pedagogique (INRP).
- Ryan, J., & Williams, J. (2007). *Children's mathematics 4-15: learning from errors and misconceptions*. Buckingham: Open University Press.